Climate Dynamics (2001) 17:213–218

G. J. Boer · S. J. Lambert

Second-order space-time climate difference statistics

Received: 26 January 2000 / Accepted: 9 June 2000

Abstract An approach to the calculation and display of second order space-time difference statistics, suitable for various applications ranging from weather forecasting to climate simulation, is discussed. The representation of the space-time agreement between model and observed quantities (or generally between any two data sets) depends on treating deterministic and random components of the variance in an appropriate way depending on context. A diagram to display the second order mean square difference, the correlation, and the ratio of variances on a single diagram in an intuitive way is also proposed. An example, comparing observed and simulated surface air temperatures from a group of models in the Coupled Model Intercomparison Program (CMIP), is presented.

1 Introduction

The comparison of fields in space and time is a common feature of meteorology and climatology. In meteorology, the comparison of forecasts with subsequent observations is used to demonstrate current levels of forecast skill, to evaluate potential changes in forecast methods, to document the improvement of forecast skill with time, and to permit the forecast user to weigh the value of the forecast in practical applications. Weather forecasts are almost universally produced with numerical models which are initialized with analyses of the atmosphere and some surface fields. The models are integrated forward in time for periods of days to weeks. The skill of an operational forecast measures the extent to which the forecast evolution of some variable correctly predicts the actual evolution of that variable in space and time. Forecast skill decreases comparatively rapidly because of the chaotic nature of the atmosphere where small errors in initial conditions and in forecast models amplify rapidly resulting in the well-known two week "predictability limit" for deterministic forecasts.

Climate forecasts, for periods beyond the deterministic predictability limit, do not attempt to predict the actual evolution of the system in detail but rather the evolution of some climate statistic. Climate forecasts made with models nevertheless follow the same general approach as weather forecasts. The climate model, which in this case may include both the atmosphere and the tropical or global ocean, is initialized based on available data and subsequently integrated forward in time. In this case, the details of the evolution of the forecast are not expected to parallel those of the observations throughout the forecast period. Certain climate statistics, such as monthly or seasonal anomalies from the climatological mean may, however, exhibit useful skill. This is measured by comparing the forecast and observed values of these statistics.

Coupled climate models attempt to simulate the observed distribution of the climatological statistics of the system which arise as a consequence of the governing physical processes and balances embodied in the model equations. A climate model is initialized with more or less arbitrary initial conditions which may, but need not, correspond to observations of the real system. In either case, the early part of a climate simulation, where the modelled system is coming into equilibrium with its forcing, is discarded. The subsequent integration does not attempt to reproduce the actual observed evolution of the system but only one of many possible evolutions given the forcing of the system. The skill of the climate model resides in its ability to reproduce the observed distribution of the set of climate statistics.

There are many ways to measure skill but we concentrate here on the calculation and display of simple second order statistics involving mean square differences, variances, and correlations.

G. J. Boer (✉) · S. J. Lambert
Canadian Centre for Climate Modelling and Analysis,
Meteorological Service of Canada, University of Victoria,

Mailing address: PO Box 1700,
Victoria, B.C. V8W 2Y2, Canada
E-mail: George.Boer@ec.gc.ca

2 Second order statistics

A straightforward comparison of two data sets appeals to second order statistics, such as mean square differences, variances, covariances, and the associated ratios of these terms including correlation. The simple geometric interpretation of the mean square difference, shown in Fig. 1, treats the two data sets as vectors $\boldsymbol{X} = \{X(\lambda_i, \varphi_j, p_k, t_l)\} = \{X_\alpha\}, \boldsymbol{Y} = \{Y(\lambda_i, \varphi_j, p_k, t_l)\} = \{Y_\alpha\}$. The product $\boldsymbol{X} \cdot \boldsymbol{Y} = |\boldsymbol{X}||\boldsymbol{Y}| \cos\varphi = \sum_\alpha w_\alpha X_\alpha Y_\alpha$ is written in terms of generalized vector lengths $|\boldsymbol{X}|$ and $|\boldsymbol{Y}|$ and a measure, $\cos\varphi$, of the angle between the vectors. Squared vector lengths follow as $\boldsymbol{X} \cdot \boldsymbol{X} = |\boldsymbol{X}|^2 = \sum_\alpha w_\alpha X_\alpha^2$ and similarly for $\boldsymbol{Y}$. The weights w_α are chosen to give the appropriate area and time average over the region and period of interest.

In vector notation the squared length of the difference vector, $\boldsymbol{d} = \boldsymbol{Y} - \boldsymbol{X}$, is written as

$$|\boldsymbol{d}|^2 = |\boldsymbol{Y} - \boldsymbol{X}|^2 = |\boldsymbol{X}|^2 + |\boldsymbol{Y}|^2 - 2\boldsymbol{X} \cdot \boldsymbol{Y}$$
$$= |\boldsymbol{X}|^2 + |\boldsymbol{Y}|^2 - 2|\boldsymbol{X}||\boldsymbol{Y}| \cos\varphi \tag{1}$$

and in scalar notation as

$$\overline{d^2} = \overline{(Y - X)^2} = \sigma_X^2 + \sigma_Y^2 - 2\sigma_X\sigma_Y r \ . \tag{2}$$

The overbar represents a generalized averaging operation in space and time corresponding to the weighted sum in the expression for the vector dot product, the lengths of the vectors are represented by σ_X and σ_Y, and $r = \cos\varphi = \boldsymbol{X} \cdot \boldsymbol{Y}/|\boldsymbol{X}||\boldsymbol{Y}|$ is a measure of the angle between the vectors. We generally adopt the "centred" approach where the means of each data set are subtracted out before the terms in Eq. (2) are calculated so that σ_X and σ_Y are the usual standard deviations (their squares the variances), and r the centred correlation coefficient. However, the results follow also in the "uncentred" case when root mean squares and uncentred correlations are used.

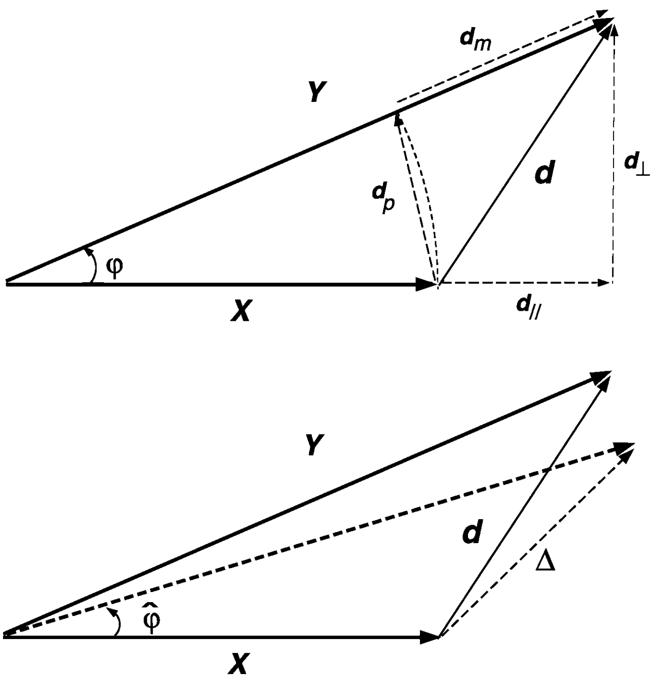

Fig. 1 Graphical representation of the mean squared difference between two data vectors. The length of the data vectors is measured by the standard deviation, and the length of the difference vector by the root mean squared difference. The cosine of the angle between the vectors represents the correlation. The difference vector may be decomposed into components in several ways. The *climate mean squared difference* takes into account the difference between deterministic and random components resulting in a reduction of the angle between the vectors and a reduced difference vector as indicated in the *bottom panel* and explained in the text

2.1 Mean square difference components

The geometric view and Eqs. (1, 2) illustrate the link between the mean square difference (*msd*) and the correlation, the two common second order measures of the agreement between fields. The geometric view suggests the further decomposition of the difference vector as $\boldsymbol{d} = \boldsymbol{d}_\parallel + \boldsymbol{d}_\perp$ with components respectively in the direction of and perpendicular to the reference vector **X** where

$$|\boldsymbol{d}|^2 = |\boldsymbol{d}_\parallel|^2 + |\boldsymbol{d}_\perp|^2 = \overline{d^2} = (\sigma_Y r - \sigma_X)^2 + \sigma_Y^2(1 - r^2) \ . \tag{3}$$

The difference vector may also be written as $\boldsymbol{d} = \boldsymbol{d}_m + \boldsymbol{d}_p$, as illustrated in Fig. 1, where, because the components are not orthogonal,

$$|\boldsymbol{d}|^2 = |\boldsymbol{d}_m|^2 + |\boldsymbol{d}_p|^2 + 2\boldsymbol{d}_m \cdot \boldsymbol{d}_p = \overline{d^2}$$
$$= (\sigma_Y - \sigma_X)^2 + 2\sigma_X^2(1 - r) + 2\sigma_X(\sigma_Y - \sigma_X)(1 - r) \tag{4}$$

Here $|\boldsymbol{d}_m|^2 = (\sigma_Y - \sigma_X)^2$ is the mean square difference that arises because of differences in the lengths of the data vectors but not their orientation (as would be the case if $r = 1$). Similarly, $|\boldsymbol{d}_p|^2 = 2\sigma_X^2(1 - r)$ is the *msd* that arises from the difference in orientation of the vectors (a difference of pattern in the sense of correlation) but not of length (as would be the case if $\sigma_Y = \sigma_X$). The remaining contribution $2\boldsymbol{d}_m \cdot \boldsymbol{d}_p = 2\sigma_X(\sigma_Y - \sigma_X)(1 - r)$ is the consequence of there being differences in both magnitude and pattern (length and orientation of the vectors) simultaneously.

The *msd* may also be written in a form with a less obvious geometric interpretation as

$$\overline{d^2} = \overline{(Y - X)^2} = (\sigma_X^2 + \sigma_Y^2)(1 - \beta r) = (\sigma_X^2 + \sigma_Y^2)(1 - p) \tag{5}$$

where $\beta = 2\sigma_X\sigma_Y/(\sigma_X^2 + \sigma_Y^2) \leq 1$ and $p = \beta r$ is a scaled correlation coefficient. Here $\sigma_X^2 + \sigma_Y^2$ is the *msd* when the data vectors are orthogonal, i.e. when the correlation between the fields is zero. The pattern factor $1 - p$ gives the fraction of this limiting value.

2.2 Scaled statistics

The *msd* is often scaled to give the relative or fractional value $f = \overline{d^2}/\sigma_X^2$ which measures the *msd* in terms of the reference variance and this is done here. In terms of Fig. 1, the vectors are rescaled so that the reference data vector has unit length. Other scalings may be appropriate depending on the purpose, for instance Boer (1994) scales the *msd* between weather forecasts and observations as $f = \overline{d^2}/(\sigma_X^2 + \sigma_Y^2) = 1 - \beta r$ where the decorrelation of forecast and observations with forecast range gives the limiting values $f \to 1$ as $r \to 0$.

3 Deterministic and random components

There are many variables of interest for weather forecasting and climate modelling but attention is usually concentrated on relatively few of them. For instance, 500 mb height is a standard forecast variable and the skill at forecasting this variable is often intercompared by operational forecast centres. Temperature and precipitation are particularly important for forecasts at all time ranges and are the variables usually treated in seasonal and multi-seasonal forecasts (e.g. Kirtman 2000). These variables are also basic to climate studies and to climate change simulations (e.g. Lambert and Boer 2000; Boer et al. 2000a, b).

The useful comparison of fields in both space and time using second order statistics depends on the ability to separate these fields into what are here termed *deterministic* and *random* components. The variability of some quantity may be deterministic in one case and random in another depending on the context and purpose of the comparison. As an example consider the variation of temperature at a particular point. The *msd* between the observed and forecast temperature at that point will be a measure of the deterministic forecast skill for a short-range weather forecast. For a climate simulation, the model is expected to reproduce the observed distribution of mean temperature and of the mean diurnal and seasonal cycles of temperature since these are forced deterministic components. The remaining variability is an internally generated non-linear natural random component and modelled and observed values of this component (or those from different realizations with the same model) are not expected to be correlated. The skill of a climate simulation is measured by comparing the variances and other second order statistics of simulated and observed temperatures, but not the detailed evolution implied by the temporal correlation.

We deal here primarily with climate simulations and write

$$X = X_0 + X_f + X' \tag{6}$$

where the variable X is taken as the sum of a long-term climate mean X_0, a function of space but not of time, other forced components X_f, which may be a function of both space and time, and the remaining random natural variability of the non-linear system represented as X'. Both the mean and any other forced components (such as the annual and diurnal cycles or forced climate change) are deterministic in that they are the physically determined response of the system to a particular forcing.

The forced components depend on the separation of the system into external or prescribed and internal or interactive components. For short-term weather forecasting, forecast skill consists in the ability to predict the detailed evolution from specific initial conditions. For seasonal forecasting, anomalous SSTs may be specified and act to force the atmospheric state. The skill of the seasonal forecast is in correctly producing the average forced response, rather than the detailed evolution of X'. For climate models, the mean, and the annual and diurnal cycles of temperature and other variables are forced by solar radiation while the natural variability is internally generated and will be different for different simulations. The skill of a climate simulation is in correctly reproducing the forced components together with pertinent statistics of the natural variability.

We assume that the mean and other forced components may be identified by averaging over a number of realizations or a sufficiently long simulation. We calculate and display simple second order space-time measures of agreement between model and observations where the forced deterministic and random natural variability components are combined in a suitable way.

4 Assessment and intercomparison of model results

Complicated numerical models are used for weather forecasting, climate forecasting, and climate simulation. The continued improvement of such models is a constant theme of model development. The large amounts of the data produced by these models, especially climate models, exceed that observed and/or analyzed. It is important to "assess" model results in order to know the extent to which they may be relied on in future forecasts of weather, short-term climate variability, and long-term climate change. However, the non-linearity and complexity of the climate system, and the models that attempt to simulate its behaviour, means that model deficiencies are not easily traced to their causes. Model intercomparison projects (MIPs) (Gates 1987; Boer et al. 1992; Gates et al. 1999; Boer 2000) attempt to gain insight into model behaviour by comparing model results among themselves as well as with observations.

The accompanying paper (Lambert and Boer 2000) is an example of one such MIP. In that paper, among other things, we represent aspects of simulated climate using basic second order measures including scaled mean square differences $\overline{d^2}/\sigma_X^2$, the ratio of variances $\sigma^2{}_Y/\sigma_X^2$, and the correlation r, all represented on a single diagram. Figure 2 is such a diagram which is an outgrowth of a diagram due to Taylor as described in a recent MIP publication (Gates et al. 1999) but with several important differences as discussed later*.

4.1 A "climate" mean square difference

We consider observed and modelled climate data for the current climate (or simulated climate data from two climate models). For concreteness, consider monthly mean temperatures. The data are of the form Eq. (6) where the climate mean is well estimated as the long-time mean, the forced component consists of the annual cycle about this mean, and what remains is the natural variability. The *msd* between model result Y and observations X is

$$\begin{aligned}\overline{d^2} &= \overline{(Y-X)^2} = \overline{(Y_0-X_0)^2} + \overline{(Y_f-X_f)^2} \\ &\quad + \overline{(Y'-X')^2} = \overline{d_0^2} + \overline{d_f^2} + \overline{d'^2}\end{aligned} \tag{7}$$

where the overbar is a general space-time average. We expect a climate model to be able to simulate the spatial distribution of the mean climate and the spatial distribution and temporal behaviour of the annual cycle. These are forced deterministic climate components and the *msd* between modelled and observed values is measured by $\overline{d_0^2}$ and $\overline{d_f^2}$.

This is not the case for the remaining natural variability since the climate model is not expected to

*Taylor, K.E. summarizing multiple aspects of mode performance in a single diagram. Submitted to J. Geophys. Res.

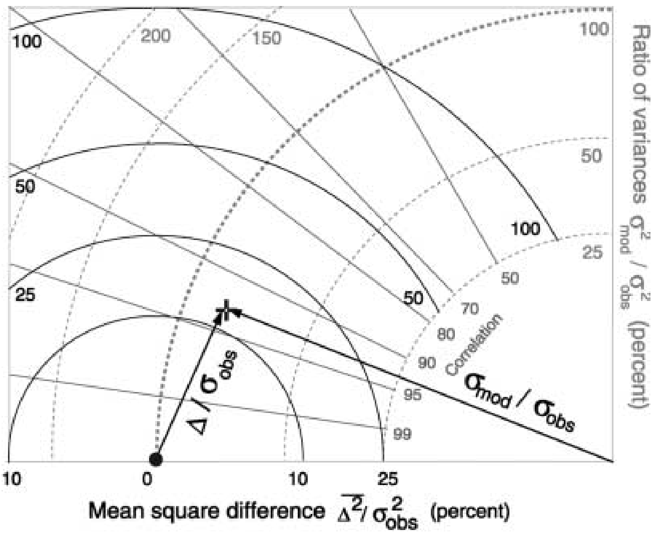

Fig. 2 The BLT diagram representing the relative climate mean squared difference, the ratio of the variances, and the effective correlation between data sets

reproduce a particular observed evolution of this component. In other words, the correlation of the modelled and observed natural variability is expected to be zero, even if the model is a perfect climate model. We do expect the model to produce the correct magnitude of the variance of the natural variability, however, and this is measured by a modified *msd* where $\overline{d'^2}$ is replaced by

$$\overline{\delta'^2} = (\sigma_{Y'} - \sigma_{X'})^2 \ . \tag{8}$$

This term involves only the difference in the observed and modelled standard deviations of the natural variability, as discussed in connection with Eq. (4), and discounts that part of the *msd* that implicitly assumes that the observed and modelled natural variability should be correlated. The modified *msd* is smaller than the usual *msd* since (for $r' = 0$ as expected),

$$\overline{\delta'^2} = \overline{d'^2} - 2\sigma_{X'}\sigma_{Y'}(1 - r') = \overline{d'^2} - 2\sigma_{X'}\sigma_{Y'} \ . \tag{9}$$

A suitable space-time *climate mean square difference*, (cmsd) is given by

$$\overline{\Delta^2} = \overline{d_0^2} + \overline{d_f^2} + \overline{\delta'^2} \ . \tag{10}$$

This can be rewritten as

$$\overline{\Delta^2} = \sigma_X^2 + \sigma_Y^2 - 2\sigma_X\sigma_Y\hat{r} \tag{11}$$

which has the basic form of Eq. (2) but with a modified correlation measure given by

$$\begin{aligned}\hat{r} = \cos\hat{\varphi} &= r + \frac{\sigma_{X'}\sigma_{Y'}}{\sigma_X\sigma_Y}(1 - r') \\ &= r + \frac{\sigma_{X'}\sigma_{Y'}}{\sigma_X\sigma_Y} > r = \cos\varphi\end{aligned} \tag{12}$$

so that $\hat{\varphi} < \varphi$. In terms of Fig. 1, the model data vector has been rotated toward the observed data vector without change of length giving an appropriate measure of the space-time difference for climate data as Δ. The geometric representation implies that $|\hat{r}| \leq 1$ which may also be shown algebraically.

This approach differs, at least in flavour, from that of Taylor as reported in Gates et al. (1999). In the limiting case of no sampling error, a "perfect" climate model would, among other things, simulate forced components that match those observed so that $\overline{d_0^2} = \overline{d_f^2} = 0$. The unforced natural variability simulated by the model would be of the same magnitude as that observed so that $\sigma_{X'} = \sigma_{Y'}$, but the simulated and observed variability would not be expected to be temporally correlated so that $r' = 0$. In this case the *msd* for a perfect climate model has a non-zero value

$$\begin{aligned}\overline{d^2} &= \overline{d_0^2} + \overline{d_f^2} + \overline{d'^2} \to \overline{d'^2} \\ &= \sigma_{X'}^2 + \sigma_{Y'}^2 + 2\sigma_{X'}\sigma_{Y'}r' \to 2\sigma_{X'}^2 > 0 \ .\end{aligned}$$

The *cmsd* is modified to adjust for the expected lack of temporal correlation between observed and modelled natural variability so that, for this measure,

$$\overline{\Delta^2} = \overline{d_0^2} + \overline{d_f^2} + \overline{\delta'^2} \to \overline{\delta'^2} = (\sigma_{Y'} - \sigma_{X'})^2 \to 0$$

in the limiting case.

4.2 The mean model *cmsd*

If simulations by different climate models are considered to be independent estimates or realizations of the climate, each with different errors, then the ensemble mean obtained by averaging over the collection of model results may at least partially average out these errors to give an improved estimate of the climate (see also the discussion in Lambert and Boer 2000). The second order climate difference statistics for the resulting *mean model* may also be calculated and compared with the statistics for individual models. Climate statistics for the mean model are obtained by treating deterministic and random components separately as in the case of the *cmsd* and for the same reasons. Mean model deterministic components are estimated as $\{Y_0\}, \{Y_f\}$ where the braces represent the ensemble mean over the collection of model results. The *mean model cmsd* is

$$\overline{\nabla^2} = \overline{\{d_0\}^2} + \overline{\{d_f\}^2} + \tilde{\delta}^2 \tag{13}$$

where the first two terms are just those in Eq. (10) but with the *mean model* value appearing in place of an individual model value, i.e. $\overline{\{d_0\}^2} = \overline{(\{Y_0\} - X_0)^2}$. This implicitly assumes a single (presumably accurate) observation-based value, but if there are several equally reasonable observation-based estimates, from different reanalyses for instance, it would also be suitable to use their ensemble average as $\overline{\{d_0\}^2} = \overline{(\{Y_0\} - \{X_0\})^2}$.

The contribution from the natural variability, analogous to Eq. (8), is given by

$$\tilde{\delta}^2 = (\tilde{\sigma}_{Y'} - \sigma_{X'})^2 \tag{14}$$

where the *mean model* random variance $\tilde{\sigma}^2_{Y'} = \{\sigma^2_{Y'}\}$ is obtained by ensemble averaging the over the collection of model variance values. If an ensemble of observational values were available then $\tilde{\sigma}_{X'}$ could appear in Eq. (14) in place of $\sigma_{X'}$. For single member "ensembles" Eq. (13) reverts to Eq. (10) as would be expected.

Although discussed in the context of comparing a *mean model* result with observations, the *cmsd* in Eq. (13) is, more generally, an *ensemble mean csmd*. It may potentially be used to calculate second order difference statistics between ensemble means of collections of model results and collections of observational estimates as well as between sets of model results with different characteristics (e.g. different resolutions, different physical parameterizations) and so on.

4.3 Graphical representation

Figure 1 is the basis of a graphical representation of second order measures of differences between data sets, and Fig. 13 of Gates et al. (1999) is such a representation. We propose, in Fig. 2, a new version of the diagram where the appropriate *cmsd* is measured by $\overline{\Delta^2}$ and the effective correlation by $\hat{r}$. Figure 2, which we refer to as a BLT diagram, is a modified version of Fig. 1 (and Fig. 13 of Gates et al. 1999) where: (1) quantities are scaled by the observed variance to give relative values, (2) the *climate* mean square difference is plotted, (3) the diagram is rotated so that *cmsd* is zero at the origin and increases away from that point, and (4) the relative *cmsd*, the variance ratio, and the effective correlation are all indicated for each point on the diagram.

The diagram may be used to plot information for different variables, and to represent the total *cmsd* and its various components in a straightforward and intuitive way.

4.4 Components

The majority of quantities intercompared in CMIP are global distributions of long-term average quantities, in effect $X = X_0$ the spatially varying climate mean. In that case we deal only with the forced component and spatial averages so $\overline{d^2} = \overline{\Delta^2}$ (i.e. *msd* = *cmsd*). In the case of the global temperature distribution, for instance, the models are relatively successful in capturing the strong equator to pole temperature structure that dominates the global variance. A typical decomposition of the spatial structure of temperature, or other variable, is

$$X = \langle X \rangle + [X]^+ + X^* \tag{15}$$

where $\langle X \rangle$ is the global mean, $[X]$ the zonal (longitudinal) mean, and $X^+ = X - \langle X \rangle$, and $X^* = X - [X]$ are the respective deviations from these means. The *msd*

$$\langle d^2 \rangle = \langle d \rangle^2 + \langle [d]^{+2} \rangle + \langle d^{*2} \rangle \tag{16}$$

represents differences in global mean temperature, differences in the north-south structure of temperature, and differences in the remaining geographical temperature structure after the mean north-south structure has been subtracted out. The ability of the climate model to reproduce this latter component is a much more stringent test than the ability to reproduce the strongly forced north-south component. This is seen in Fig. 2 of Lambert and Boer (2000) for instance.

5 The space-time *cmsd* of surface air temperature

The space-time *cmsd* is an appropriate second order measure of the difference between a modelled and observed quantity in the context of climate simulation. Surface air temperature is the only quantity in the Coupled Model Intercomparison Project (CMIP1) data base which has both the temporal and spatial information needed to calculate it, however. Figure 3 plots the space-time *cmsd* calculated from Eq. (10) where the observation-based data are 40 years of monthly mean surface air temperatures from the NCEP reanalysis. The overbar represents the average over the globe and over this time period.

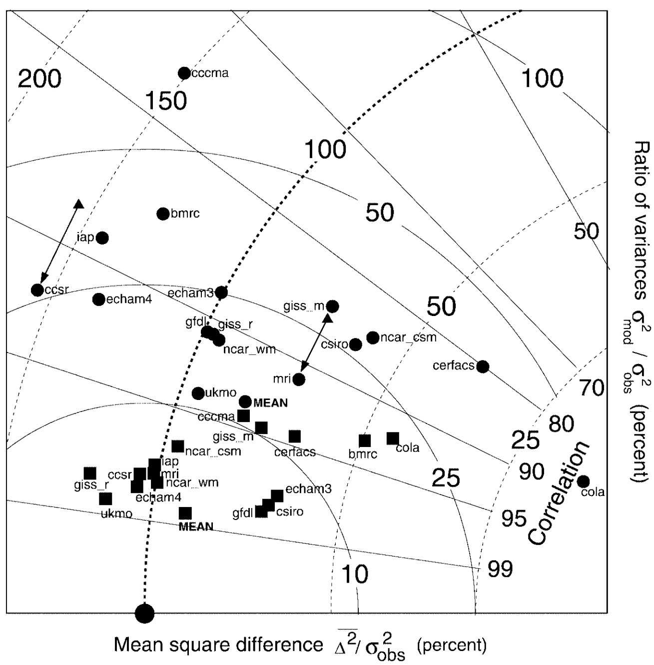

Fig. 3 The second order space-time climate mean square difference statistics for simulated and observed monthly mean surface air temperature for individual models and for the mean model. *Squares* indicate values for the total *cmsd* and *circles* for the *cmsd* of the geographic component which is obtained by removing the zonal average and hence the strong north-south structure of the temperature field. The difference between the space-time c*msd* used here and the *msd* are illustrated, in two cases, by the *arrows* from the *small triangles* (the *msd*) to the *large circles* (the *cmsd*). Simulated values are from the Coupled Model Intercomparison Project (CMIP) and observed values are from the NCEP reanalysis

The total *cmsd*, computed from the temperature field without decomposition into components, is indicated by the squares in Fig. 3 for each of 16 coupled models. The models are reasonably successful in simulating the total space-time variance of temperature (effective correlations range from 0.92 to 0.99 with a median value of 0.98). For temperature, however, a large fraction of the spatial and temporal variance is associated with the strongly forced mean north/south structure. In this circumstance, as noted in the previous section, the *cmsd* is not a very stringent test of the model's ability to simulate all features of the field in question (the *cmsd* is close to the *msd* in this case also).

We consider separately the ability of the models to simulate the geographic temperature pattern X^* which is obtained by removing the global mean and the north/south structure as indicated by Eq. (15). The difference between modelled and observed values is now $d = Y^* - X^*$ and the *cmsd* $\overline{\Delta^2} = \overline{d_0^2} + \overline{d_f^2} + \overline{\delta'^2}$ depends on differences in the mean, annual cycle, and transient variability of this component. The result is plotted in Fig. 3 as the circles. Simulating the space-time structure of this geographical component of the temperature field is a greater challenge to the models. In this case, the difference between the space-time *cmsd* and the *msd* is also larger. This is illustrated in two cases by the arrow from a small triangle (the *msd*) to a large circle (the *cmsd*) indicating the rotation of the model's data vector as described in Fig. 1.

The *cmsd* for the *mean model* is also shown in Fig. 3. Measured in this way, the mean model result is closer to the observations than most, if not all, individual models. Other mean model results are discussed in Lambert and Boer (2000).

6 Summary

A space-time climate mean square difference or *cmsd*, associated variance and correlation measures, and a diagram displaying these quantities is described for the comparison of simulated and observed fields. The *cmsd* depends on the separation of the information into deterministic forced climate components and the remaining random natural variability. The difference between the forced components is measured by the usual *msd* which depends both on differences in the variances of the fields and on their temporal and spatial correlation. The difference between natural variability components, on the other hand, is measured only by the variances since there is no physical reason for the temporal correlations to be other than zero. The combined result is a space-time climatological measure of difference suitable for comparing modelled and observed data sets or different model results.

References

Boer GJ (1994) Predictability regimes in atmospheric flow. Mon Weather Rev 122: 2285–2295

Boer GJ (2000) Climate model intercomparison. In: Numerical modelling of the global atmosphere in the climate system. Kluwer Dordrecht, The Netherlands, pp 443–464

Boer GJ, Arpe K, Blackburn M, Deque M, Gates WL, Hart TL, Le Treut H, Roeckner E, Sheinin DA, Simmonds I, Smith RNB, Tokioka T, Wetherald RT, Williamson D (1992) Some results from an intercomparison of the climates simulated by 14 atmospheric general circulation models. J Geophys Res 97: 12 771–12 786

Boer GJ, Flato G, Reader MC, Ramsden D (2000a) A transient climate change simulation with greenhouse gas and aerosol forcing: experimental design and comparison with the instrumental record for the 20th century. Clim Dyn 16: 405–425

Boer GJ, Flato G, Ramsden D (2000b) A transient climate change simulation with greenhouse gas and aerosol forcing: projected climate for the 21th century. Clim Dyn 16: 427–450

Gates WL (1987) Problems and prospects in climate modelling. In: Radok U (ed) Toward understanding climate change. Westview Press, Boulder, Colorado, pp 5–34

Gates WL, Boyle JS, Covey C, Dease CG, Doutriaux CM, Drach RS, Fiorino M, Gleckler PJ, Hnilo JJ, Marlais SM, Phillips TJ, Potter GL, Santer BD, Sperber KR, Taylor KE, Williams DN (1999) An overview of the results of the Atmospheric Model Intercomparison Project (AMIP). Bull Amre Meteor Soc 80: 29–56

Kirtman B (ed) (2000) Experimental Long-lead Forecast Bulletin. Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland. Also at www.iges.org/ellfb

Lambert SJ, Boer GJ (2000) CMIP1 evaluation and intercomparison of coupled climate models. Clim Dyn 17: 83–106